PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/24, C07K 13/00, 7/06
C12N 15/62, C12P 21/08
A61K 37/02, C07K 3/08
C12N 1/21, A61K 39/395

A1

(11) International Publication Number: WO 91/13985

(43) International Publication Date: 19 September 1991 (19.09.91)

(21) International Application Number: PCT/GB91/00334

(22) International Filing Date: 5 March 1991 (05.03.91)

(30) Priority data:
9004890.1 5 March 1990 (05.03.90) GB

(71)(72) Applicants and Inventors: HEATH, John, Kaye [GB/GB]; University of Oxford, Department of Biochemistry, South Parks Road, Oxford OX1 3QU (GB). SMITH, Austin, Gerrard [GB/GB]; Centre for Animal Genome Research, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ (GB). RATHJEN, Peter, David [AU/AU]; Department of Biochemistry, University of Adelaide, North Terrace, Adelaide, S.A. 5001 (AU).

(74) Agents: GOLDIN, Douglas, Michael et al.; J.A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5LX (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: LEUKAEMIA INHIBITORY FACTOR

(57) Abstract

A substantially pure Leukaemia Inhibitory Factor (LIF) which is capable of associating with a mammalian extracellular matrix. Preferably, the N-terminal sequence comprises: NH_2-U-arg-X-arg- where U is the residue of a peptide of 20 or fewer amino acids, or a covalent bond, and X is any amino acid. Such N-terminal sequences may also be attached to other proteins or peptides in order that they may become associated to the ECM.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

LEUKAEMIA INHIBITORY FACTOR

This invention relates to Leukaemia Inhibitory Factor, to its production and uses.

Leukaemia Inhibitory Factor (LIF) is a secreted, soluble polypeptide regulatory factor which has the characteristic of specifically suppressing the differentiation, and maintaining pluripotentiality, of murine embryo-derived pluripotential embryonic stem cells.

Experiments using ES cell lines established directly from the inner cell mass of the early mouse embryo (Evans and Kaufman, 1981; Martin, 1981), suggest that LIF is a normal embryonic regulatory factor. In the presence of LIF, also known as DIA (differentiation inhibiting activity), ES cells can be cultured and maintained *in vitro* without loss pluripotentiality. Even after many generations *in vitro*, these cells can be reintroduced into the mouse blastyocyst where they may contribute differentiated progeny to all tissues including the germ line (Bradley *et al* 1984).

EP-A-285 448 discloses the complete amino acid sequence of a form of murine LIF and also discloses a sequence, incomplete at the N-terminus, of the humar form of this factor. The murine LIF disclosed by EP-A-285 448 has the N-terminal sequence NH_2-met-lys-val-leu-ala-ala-gly (KVL-LIF).

Stahl et al (J.Biol Chem., 265(15),8833 (May 1990))

disclose the genomic sequence of human LIF.

We have now surprisingly found that certain cell lines which exhibit DIA-LIF activity do so via a novel form of LIF which has the N-terminal sequence NH_2-met-arg-cys-arg (RCR-LIF). This form of LIF occurs as a result of a splicing event which joins a new 5'exon to the second and third exons of the previously described form of LIF. This alternative splicing thus replaces the first seven amino acid residues of KVL-LIF with the first four amino acids of RCR-LIF. The result of this alteration in the amino terminus of the LIF protein results in the incorporation of RCR-LIF into the extracellular matrix (ECM). RCR-LIF is therefore a localised, as opposed to diffusible, determinant of differentiation.

As a result of our findings, we have identified two proteins, laminin-β2 and e-cadherin, which form part of the normal ECM which have, at or near their N-termini, the sequence:

-arg-J-arg-

where J is glycine in laminin-β2 and cysteine in e-cadherin. However, since these proteins share many other sequence similarities, it was not previously apparent whether any part of these proteins - and if so, which - was responsible for their presence in the ECM.

We have found that the localization of RCR-LIF can be demonstrated by the inability of this form of LIF, when

associated with an ECM, to diffuse through a microporous membrane. As demonstrated in the following examples, a cell line responsive to the action of LIF, such as a pluripotent cell line (for example the ES cell line mentioned above), will only respond to an ECM containing this form of LIF when in direct contact with the ECM. The presence of a microporous membrane between an ECM containing RCR-LIF and LIF-responsive cells will block the action of RCR-LIF on the cells, but not the action of the diffusable KVL-LIF, which is not retained by the ECM. The RCR-LIF thus remains associated with the ECM. In the absence of a physical barrier such as a microporous membrane, both forms of LIF act upon LIF-responsive cells.

RCR-LIF is thus distinguished from KVL-LIF by the following criteria:

1. LIF activity is physically confined to the extracellular matrix (ECM) and is thus only acts upon cells which are in direct physical proximity to ECM containing RCR-LIF.

2. ECM associated RCR-LIF source has superior stability to KVL-LIF.

3. RCR-LIF activity is confined to cells which are capable of physically associating with the RCR-LIF associated ECM preparation and therefore its range of biological activities in vivo and in vitro can be targeted more sp cifically than KVL-LIF.

4. RCR-LIF can be presented in a more concentrated form (as it is not freely diffusible) with improved biological efficacy compared to KVL-LIF derived from non-ECM associated sources.

In addition, RCR-LIF as defined by these parameters has the following properties:

1. Enhances the survival of sympathetic and CNS derived neurons in vivo and in vitro.

2. Regulates the expression of neurotransmitters, for example, choline acetyltransferase and acetylcholine esterase, in neuronal preparations.

3. Induces the proliferation of cells of the osteoblast lineage in vivo and in vitro.

4. Induces activation of osteclastic cells in osteogenic tissues.

5. Supresses the expression of lipoprotein lipase in cells of the adipocytic phenotype in vivo and in vitro.

6. Induces the expression of acute phase response proteins in hepatocytes or hepatocyte derived cell types.

7. Enhances the viability and multiplication of keratinocytes in vitro.

8. Maintains the proliferation of haemopoetic cells.

9. Induces the differentiation of leukaemic cells.

10. Enhances the viability and multiplication of

germ cells.

The novel N-terminal sequence of the LIF of the present invention provides a signal sequence for directing and anchoring molecules such as proteins or peptides to a mammalian ECM. Therefore, the tetrapeptide sequence NH_2-met-arg-cys-arg may be attached by chemical or recombinant means to proteins or peptides to cause the molecules to become attached to or associated with the ECM. C-terminal derivatives of this tetrapeptide sequence corresponding to further LIF sequence, eg. NH_2-met-arg-cys-arg-ile-val, may also be used. It is to be understood that changes in the tetrapeptide sequence NH_2-Met-Arg-Cys-Arg which do not substantially alter the ability of this sequence to direct association with the ECM are within the scope of the invention.

The LIF of the present invention can be used for the propagation of embryonic stem (ES) cells and for the maintenance of ES cell pluripotentiality (as defined by the ability to form functional gametes in chimeras). In addition, LIF of the invention and N-terminal fragments thereof, and polyclonal or monoclonal antibodies or fragments thereof against LIF and its N-terminal fragments have potential utility for the following applications:

1. Induction and suppression of haemopoetic cell differentiation in vivo and in vitro.

2. Induction of bone and tooth deposition and

resorption in vivo.

3. Regulation of acute phase response protein gene expression in liver cells.

4. Regulation of cachexia.

5. Regulation of neuron viability and differentiation.

6. Acceleration/Inhibition of wound healing.

7. Induction and inhibition of epithelial cell proliferation and differentiation.

The present invention thus provides a novel LIF which is capable of associating with a mammalian extracellular matrix. In one aspect of the invention, this novel LIF is characterised by an N-terminal sequence NH_2-met-arg-cys-arg.

The invention also provides novel forms of LIF containing sequences which are substantially identical to the sequence met-arg-cys-arg which are capable of causing LIF to be associated with the ECM. In particular, homologues of this LIF sequence in which 2 of the first 5 amino acid residues from the N-terminus are substituted are within the scope of the invention.

A further embodiment of the invention provides a DNA coding for any of these novel forms of LIF, N-terminal signal sequences, or recombinant proteins or peptides described above. The DNA may be used to produce a labelled probe by conventional means using radioactive or non-

radiaoactive labels, or the DNA may be cloned into a vector.

A further embodiment of the invention provides vectors for the replication and expression of the said DNA. The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said DNA and optionally a regulator of the promoter. The vector may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. The vector may be used in vitro, for example for the production of RNA corresponding to the DNA, or used to transfect or transform a host cell.

A further embodiment of the invention provides host cells transformed or transfected with the vectors for the replication and expression of the said LIF DNA. The cells will be chosen to be compatible with the vector and may for example be bacterial, yeast or mammalian.

The invention also provides N-terminally directed monoclonal or polyclonal antibodies, or fragments thereof, to the novel proteins or peptides, or N-terminal fragments thereof, of the invention.

The invention also provides a process for the production of monoclonal or polyclonal antibodies to the novel proteins or peptides of the invention. Monoclonal

antibodies may be prepared by conventional hybridoma technology using the novel proteins or peptides, or fragments thereof, as an immunogen. Polyclonal antibodies may also be prepared by conventional means which comprise innoculating a host animal, for example a rat or a rabbit, with a protein or peptide of the invention and recovering immune serum.

The present invention also provides pharmaceutical compositions containing the LIF proteins or peptides of the invention, or fragments thereof, and antibodies or fragments thereof to the proteins, peptides and fragments thereof for the treatment and/or regulation of conditions, including proliferative diseases, associated with abnormal (eg at an unusually high or low level) and/or aberrant (eg due to a mutation in the gene sequence) expression of LIF in a mammal, including man, or for the treatment of conditions or diseases of a mammal, including man which benefit from an alteration in naturally occuring levels of LIF.

The invention also provides the above protein, peptides and antibodies and fragments thereof, and DNA coding for the said proteins and peptides (or fragments thereof) for the therapy or diagnosis of conditions, including proliferative diseases, associated with abnormal and/or aberrant expression of LIF in a mammal, including man.

In addition, the invention provides peptides of the sequence NH_2-met-arg-cys-arg, C-terminal derivatives thereof, eg NH_2-met-arg-cys-arg-ile-val, and substantially identical sequences as defined above for use as N-terminal signal sequences for naturally occurring molecules lacking such a sequence or any form of recombinant molecules. Such sequences will cause such molecules, eg proteins or peptides, to become associated with the ECM.

Thus, the present invention thus further provides recombinant protein or peptide which comprises, at its N-terminus, the sequence:

NH_2-U-arg-X-arg-

where U is the residue of a peptide of 20 or fewer amino acids, or a covalent bond, and X is any amino acid.

Preferably, U is 20 or fewer, eg 10 or fewer residues in size. Preferably, X is glycine or cysteine. Preferably, the N terminal sequence comprises:

NH_2-U-arg-X-arg-ile-val-pro-; or

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U and X are as defined above, and substantially identical variants thereof. It will be understood that substantially identical sequences are those which do not substantially alter the ability of the N-terminal sequence to direct the protein or peptide to which the sequence is attacted to the ECM.

The protein or peptide to which the N-terminal

sequences of the invention may be attached may, for example, be a hormone, growth factor or cytokine. Attachment of sequences according to the invention to such proteins and peptides will cause the resulting recombinant proteins and peptides to become attached to the ECM. This will enable the action of these proteins and peptides to be localized. It will be appreciated by those of skill in the art that, for example, a recombinant hormone which comprises an N-terminal sequence according to the invention together with all, or an active fragment of, the hormone, will still be considered a hormone for the purposes of the invention. Similarly, other naturally occuring proteins or peptides modified according to the invention will be considered equivalent to the native protein or peptide if their function is substantially unaltered.

Such modified proteins or peptides may be used, where appropriate, in the treatment or therapy of the human or animal body.

The invention further provides recombinant DNA encoding a protein or peptide according to the invention.

The invention also provides a recombinant replicable expression vector comprising such DNA, a host cell transformed or transfected with such a vector, and a polyclonal or monoclonal antibody directed against the N-terminal of a protein or peptide according to the invention. Reference may be made to the description above

in connection with LIF for suitable vectors, host cells and processes for the production of antibodies.

The present invention also provides a peptide of the formula:

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U is the residue of a peptide of 20 or fewer amino acids or a covalent bond and X is any amino acid, and substantially identical variants thereof.

In addition, the invention provides a method of causing a protein or peptide to become attached to the ECM of a mammalian cell comprising contacting the ECM with a protein or peptide having an N-terminus of the sequence:

NH_2-U-arg-X-arg-; or

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U is the residue of a peptide of 20 or fewer amino acids or a covalent bond and X is any amino acid, and substantially identical variants thereof.

Preferably, the variants of all the sequences mentioned above will retain the motif "-arg-X-arg", where X is as defined above.

The following examples describe the isolation and characterization of the novel ECM-associated LIF of the invention from a murine source. However, ECM-associated LIF from other sources, e.g. human or porcine are within the scope of the present invention and could be isolated in an analogous manner.

EXAMPLE 1: MATERIALS & METHODS

CELL CULTURE AND BIOLOGICAL ASSAYS Cell culture was performed in a humidified 7.5% CO_2 atmosphere in DMEM:Ham's F12 (50:50) containing antibiotics, 10^{-4}M 2-mercaptoethanol and 10% (v/v) foetal bovine serum (selected batches). For maintainance of ES cells, medium was further supplemented with a 1/1000 dilution of supernatant conditioned by exposure to Cos-7 cells transfected with the pC10-6R human DIA/LIF expression plasmid (Smith et al., 1988). Ehrlich ascites cells and MRC-5 human embryonic lung fibroblasts were obtained from the Sir William Dunn school of Pathology Cell Bank, University of Oxford. BRL and STO cells (Smith and Hooper, 1987) and C3H 10T1/2 mouse embryo fibroblasts (Edwards et al., 1987) were as previously described. Differentiation assays were performed on the pluripotential ES cell line CP1, generously provided by Dr. Martin Evans, Department of Genetics, University of Cambridge. Diffusible DIA/LIF was assayed by plating ES cells at an initial density of 10^4 cells/ml/16mm well and culturing for 4 days in the presence of experimental samples. The samples were then fixed and stained with Leishman's and inhibition of morphological differentiation was assessed by microscopic inspection (Smith and Hooper, 1987).

Feeder layers were prepared by treatment of confluent cultures with mitomycin C (10μg/ml) for 2-3

hours, followed by trypsinisation and replating (Martin and Evans, 1975). The collagenous membrane inserts (Transwell^{TM}, Costar) were transparent with a pore size of 0.4μm.

Cell-free ECM preparations were obtained by hypotonic lysis with 0.02M NH_4OH (Gospodarowicz, 1984) or by detachment of cells with 0.02% EDTA in PBS (Rheinwald and Green, 1975).

DNA MANIPULATIONS. Standard techniques (Maniatis et al.) were used for the construction of all plasmids. The LIF numbering system used throughout this paper refers to the partial length murine cDNA described by Gearing et al. (1987).

A cDNA encompassing the entire murine LIF coding region (Gearing et al., 1988) was obtained by the polymerase chain reaction (PCR) on first strand cDNA generated from Ehrlich ascites RNA from Ehrlich ascites cells. Oligo dT was used to prime reverse transcription of 2.5μg of poly A^+ RNA (cDNA Synthesis Kit, Amersham). After phenol/ chloroform extraction, RNA was hydrolysed by incubation at 4° C overnight in 0.25M NaOH, the solution was neutralised by the addition of HCl, and the cDNA was purified by Sephadex G-50 chromatography and ethanol precipitation. 20ng of the cDNA was used as a template for PCR amplification using the mLIF specific primers LIFI

5'...AAGAATTCCATAATGAAGGTCTTGG...3' and LIF2 5'...AAGAATTCAGTCCATGGTACATTCAAGA...3'. These primers, which both contain EcoRI restriction sites, hybridise to the mLIF cDNA at residues 11 - 28 and 657 - 636 respectively. PCR conditions were as recommended by Perkin Elmer-Cetus. Amplification proceeded through a cycle of denaturation for 2 minutes at 94° C, annealing for 2 minutes at 60° C and polymerisation for 3 minutes at 72° C. Amplified DNA was digested with EcoRI, purified from agarose gels using Geneclean (BIO 101, San Diego, CA) and cloned into the EcoRI restriction site of pBluescript II KS+ (pDR1) and the expression vector pXMT2 (pDR10).

A probe specific for the 5' end of the mouse LIF open reading frame was constructed by digestion of pDR1 with SmaI and religation to give pDR2 which lacks mLIF sequences downstream of the mLIF SmaI restriction site.

A cDNA encoding the matrix-associated form of DIA/LIF was constructed in the expression vector pXMT2. DNA derived by RACE PCR cloning (see below) was removed from pDR101 by XhoI/SmaI digestion and cloned into XhoI/SmaI cut pDR10 to produce the plasmid pDR11.

RNA ANALYSIS. High specific activity riboprobes were synthesised by <u>in vitro</u> transcription of murine LIF cDNA fragments cloned into pBluescript II KS+ (Stratagene). Antisense probes were generated by transcription with T7

polymerase (New England Biolabs) after linearisation of the template by digestion with HindIII (pDR1, pDR2) or XhoI (pDR100, pDR101). A riboprobe specific for the 3' end of the mLIF open reading frame was produced by transcription of pDR1 with T7 RNA polymerase after linearisation of the plasmid by digestion with DdeI which cuts at nucleotide 353 in the mLIF cDNA. The reaction conditions were as described by Krieg and Melton (1988) except that 6.25µl of α-^{32}P-UTP (800 Ci/mmol, 40mCi/ml, Amersham) was used in a 15µl reaction volume.

Ribonuclease protection assays were carried out essentially as described by Krieg and Melton (1988). 15µg of cytoplasmic RNA (Edwards et al, 1985) was hybridised with 6 x 10^4 cpm probe (specific activity 3 x 10^8 cpm/µg) in the absence of added tRNA. Hybridisations were carried out at 45°C for 16-20 hours. RNA hybrids were digested with 40 µg/ml RNase A (Boehringer) and 2 µg/ml RNase T1 (BRL) at 4°C for 30 minutes. Protected fragments were analysed on a 5% polyacrylamide/ 8M urea sequencing gel which was dried before autoradiography.

RACE cDNA CLONING. 10µg of cytoplasmic RNA from Ehrlich ascites cells was reverse transcribed from the oligonucleotide 5'...ACACGGTACTTGTTGCA...3' which hybridises to residues 500 - 484 of the murine LIF cDNA. The RNA was added to 20pmol of primer and 2µl of 10 x

anneal buffer (500mM Tris/HCl pH8.0, 60mM $MgCl_2$, 400mM KCl) in a total volume of 16µl, heated to 85°C for 5 minutes and cooled slowly to room temperature. The elongation reaction was carried out as described by Frohman et al. (1988). Excess oligonucleotide was removed by gel filtration through a 2ml Sephacryl S-400 (Pharmacia) column equilibrated with 0.05xTE. 50µl fractions corresponding to the cDNA radioactive peak were pooled, concentrated by vacuum centrifugation and resuspended in 23µl of H_2O. To tail the 3' end of the cDNA with dG residues 3µl of 10mM dGTP, 6µl of 5 x tailing buffer (BRL) and 15 units of terminal deoxynucleotidyl transferase (BRL) were added and the mixture was incubated at 37°C for 60 minutes and then 70°C for 15 minutes. After ethanol precipitation the cDNA template was resuspended in 500µl of H_2O.

PCR was carried out using a mouse LIF specific oligonucleotide 5'...TTCTGGTCCCGGGTGATATTGGTCA...3' (mLIF residues 389 - 365) and an anchor oliognucleotide 5'...CCATGGCCTCGAGGGCCCCCCCCCCCCC...3'. The anchor oligonucleotide includes the restriction sites ApaI, NcoI, XhoI and SfiI while the mLIF oliogonucleotide includes the unique SmaI restriction site at nucleotide 379 in the mLIF open reading frame. PCR was carried out in a final volume of 50µl containing 7µl of the cDNA template and 25 pmol of each oligonucleotide. Reaction conditions were as recommended by Perkin Elmer-Cetus except that a final

c ncentration of 3mM $MgCl_2$ was used. DNA was denatured prior to the addition of Taq polymerase (Perkin Elmer-Cetus) by heating the reaction mixture to 94°C for 5 minutes. Each PCR cycle consisted of denaturation for 2 minutes at 94°C, annealing for 2 minutes at 55°C and elongation for 3 minutes at 72°C. Products of this reaction were analysed by Southern blot hybridisation using probes derived from pDR1, pDR2 or the LIFI oligonucleotide. After the final elongation (30 minutes at 72°C) samples were ethanol precipitated, digested with SmaI and XhoI, and analysed by agarose gel electrophoresis. DNA was purified from agarose gels using Geneclean and cloned into XhoI/SmaI digested pBluescript II KS+ (Stratagene). Suitable recombinant plasmids were purified by the alkaline lysis method. The clone corresponding to the diffusible form of DIA/LIF was designated pDR100 and that corresponding to the matrix-associated form of DIA/LIF was designated pDR101.

To sequence the RACE cDNA clones, PDR100 and pDR101 were digested with XhoI and SmlI and the DIA/LIF cDNA fragments were purified from agarose gels using Geneclean and cloned into SalI/SmaI digested M13mp18 and M13mp19. Plasmids were sequenced with Sequenase Version 2.0 (United States Biochemicals, Cleveland) according to the manufacturers recommendations.

EXAMPLE 1 RESULTS

IDENTIFICATION OF MATRIX-ASSOCIATED DIA/LIF ACTIVITY PRODUCED BY C3H 10T1/2 EMBRYONIC FIBROBLASTS

The maintainance of ES cells in an undifferentiated state can be achieved by growing the cells in medium conditioned by preincubation with BRL cells. This results from secretion of soluble DIA/LIF protein into the medium by BRL cells. By contrast, medium conditioned by different feeder cell lines may be unable to prevent or be relatively inefficient at preventing ES cell differentiation. This suggests that at least part of the DIA/LIF activity associated with feeder cells is not diffusible but is localised to the cell in some way. This could involve attachment to either the feeder cell membrane or extracellular matrix.

The results of an experiment designed to test this are illustrated in Figure 1. ES cells cultured in direct contact with 10T1/2 embryonic fibroblast cells were maintained in an undifferentiated state (Figure 1, 1). This confirms that 10T1/2 cells, in common with other embryonic fibroblast lines such as STO and 3T3, can act as feeder cells for the culture of undifferentiated ES cells. No significant DIA/LIF activity could be detected in medium that had been conditioned by preincubation with 10T1/2 cells (Figure 1. ii). Furthermore, physical separation of the 10T1/2 feeder cells and ES cells by a microporous

membrane, which permits the passage of diffusible macromolecules but prevents direct cell-cell contact, resulted in differentiation of the ES cells (Figure 1, iii). The requirement for physical contact between 10T1/2 feeder and ES cells and the absence of any secreted, diffusible DIA/LIF activity argues for the existence of a localised form of the DIA/LIF protein on 10T1/2 cells.

To identify the location of this activity 10T1/2 cells were removed from the culture by treatment with either ammonium hydroxide, which causes osmotic lysis, or EDTA, which results in the detachment of intact cells from the extracellular matrix. Since cells grow above the secreted extracellular matrix in vitro, the effect of these treatments was to yield a substrate composed almost exclusively of extracellular matrix. ES cells seeded onto this substrate were maintained in an undifferentiated state (Figure 1, iv, v) as judged by their overt morphology or the presence of the SSEA-1 antigen.

This result established the existence of a matrix-associated DIA/LIF activity on 10T1/2 fibroblasts. Identical experiments failed to detect an equivalent activity on the extracellular matrix of MRC-5 human fibroblast cells which secrete the soluble form of DIA/LIF into the medium. The matrix-associated DIA/LIF activity of 10T1/2 cells must therefore be a specific activity produced by these cells and not a non-specific effect of the

extracellular matrix. Preliminary purification and acid stability characteristics suggest that this activity is very closely related to the known DIA/LIF glycoprotein.

IDENTIFICATION OF A TRANSCRIPT ENCODING THE MATRIX ASSOCIATED FORM OF DIA/LIF

The probe used for Northern blots and the initial ribonuclease protections was derived from a cDNA fragment that contained the entire murine LIF (mLIF) open reading frame as defined by Gearing et al. (1987). This cDNA was obtained by PCR on RNA from Ehrlich ascites cells which had previously been shown to secrete soluble DIA/LIF. To test for biological activity the mLIF fragment was cloned into the unique EcoRI restriction site of pXMT2, an expression vector which contains the SV40 origin of replication. Transfection of this molecule, pDR10, into cos7 cells and bioassay of conditioned medium by its ability to prevent ES cell differentiation established that the biological activity of this clone was found to be equivalent to that of the human LIF gene described by Moreau et al. (1988)

The mLIF cDNA was also cloned into the unique EcoRI restriction site of pBluescript II KS to give pDR1. This plasmid was transcribed in vitro with T7 RNA polymerase in the presence of α-^{32}P-UTP was used to produce radiolabelled antisense RNA probes of high specific activity for use in both Northern blots and ribonuclease protections.

Northern blot analysis of total and poly /A+ RNA from Ehrich ascites (EA) and BRL cells revealed a single transcript in each cell type which varied in size from 4kb in the mouse cells to 4.5kb in the Buffalo rat cells (Figure 2). We have been unable to detect alternative LIF transcripts of 1.8kb (Moreau et al 1988) or 0.7kb (Gough et al 1988) by Northern blot despite having screened a variety of cell line and embryonic RNAs.

Total RNA from Ehrlich ascites cells was also screened for DIA/LIF expression by ribonuclease protection. Two transcripts were detected by this technique (Figure 3A). The similar sizes of the two protected bands argues that they are identical throughout the majority of the mLIF coding region since the ribonuclease protection technique identifies differences between probe and substrate by cleavage at the region of dissimilarity. To determine whether the divergence between these two messages was at the 5' or the 3' end of the mLIF coding region, riboprobes specific for either end of the open reading frame were produced. Of these, only the 5' probe protected two different transcripts (Figure 3B). This establishes that the site of divergence between the transcripts occurs at the 5' end of the mLIF open reading frame, 20+/-5 nucleotides into the coding region. This coincides with the beginning of the second mLIF exon as deduced from the genomic human LIF sequence and the partial genomic mLIF

sequence.

The identity of the two transcripts was established by analysis of DIA/LIF transcription in two embryonic fibroblast feeder cell lines, 10T1/2 and STO, as well as in EA and PYS-2 (a parietal yolk sac cell line) (Figure 4A). The upper transcript, D, corresponds to the secreted and diffusible form of the DIA/LIF protein since it shares complete homology with the riboprobe throughout the LIF coding region. This identification is consistent with the presence of significant levels of this transcript in RNA from Ehrlich ascites and STO cells which secrete DIA/LIF activity but not in 10T1/2 cells which do not secrete DIA/LIF activity. The identity of the lower transcript, M, could be deduced from a comparison of the ribonuclease protection pattern and known biological properties of the cell lines (Figure 4B). The appearance of this transcript in both STO and 10T1/2 cells suggests that it encodes the matrix-associated form of the DIA/LIF protein which allows these cell lines to act as ES cell feeders. This correlation is most apparent in 10T1/2 cells where only the lower transcript is present at significant levels and where only the matrix-associated form of the protein can be detected.

These results suggested that the production of matrix-associated and diffusible forms of the DIA/LIF protein is governed by alternative splicing of upstream

exons to the known second and third exons of the mLIF gene.

RACE PCR CLONING OF cDNA ENCODING THE MATRIX-ASSOCIATED FORM OF DIA/LIF

The race PCR technique was used to clone the unknown 5' sequences of the mRNA encoding the matrix-associated form of DIA/LIF. This technique can be used whenever sequence within the open reading frame is known and can be modified to clone sequences either 5' or 3' of the known sequence. Ribonuclease protections had shown that the matrix-associated form of DIA/LIF is identical to the soluble form of the protein throughout exons 2 and 3. An oligonucleotide complementary to residues 500 - 484 of the mLIF cDNA (Gearing et al., 1988) was used to prime cDNA synthesis from Ehrlich ascites total RNA. The cDNA was tailed with dGTP residues at the 3' end using terminal deoxynucleotide transferase and the resultant single stranded DNA was used as a template for PCR using an oligonucleotide complementary to residues 389-365 of the mLIF sequence and an anchor oligonucleotide, 5'...CCATGGCCTCGAGGGCCCCCCCCC...3'. The internal oligonucleotide was chosen both to increase the specificity of the PCR reaction and because it incorporated the unique SmaI restriction site of the mLIF open reading frame. This was used later in reconstruction of the complete open reading frame.

THE PCR did not yield the expected two bands but rather a smear of material between 400 and 600 nucleotides which was homologous to mLIF as ascertained by Southern blot. The products of an entire PCR reaction were digested with XhoI and SmaI and separated on a 2% agarose gel. Zones of DNA corresponding to mLIF sequences were excised from the gel and cloned into XhoI/SmaI cut pBluescript II KS+S. Recombinant colonies were screened for the size of the inserted fragment and the largest inserts among two distinct size classes were subcloned into M13 and sequenced. The sequences of these two clones upstream of residue 37 are shown in Figure 5 with the deduced amino acid sequences of the respective proteins. The larger cDNA (pDR100) encodes the diffusible form of the DIA/LIF protein. The sequence shown in Figure 5 is identical to that of Gearing et al (1988) but extends 122 nucleotides upstream of the ATG initiation codon. This may represent the normal 5' end of the mRNA encoding the diffusible form of the DIA/LIF protein.

The smaller clone (pDR101) diverges from the known mLIF sequence at precisely the point deduced to be the start of the second exon and is therefore a good candidate to encode the matrix-associated form of the DIA/LIF protein. The amino acid sequence of the protein encoded by this cDNA differs from that of the diffusible from of DIA/LIF in that the first part of the leader sequence,

MKVLAAG, is replaced with the sequence MRCR. The identity of the ATG initiation codon is confirmed by the presence of an in-frame termination codon, UAG, 30 nucleotides upstream of the ATG. The matrix-associated protein retains the signal sequence of the diffusible mLIF protein between amino acid 8 and the cleavage site after residue 23. This leader sequence contains the hydrophobic core sequence that is required for translocation of the LIF proteins across the cell membrane. The sequence of the pDR101 cDNA extends 31 nucleotides upstream of the ATG. This may represent the 5' end of the RNA encoding the matrix-associated DIA/LIF protein.

To confirm the identities of the two RACE clones riboprobes derived from pDR100 and pDR101 were used to protect cytoplasmic RNA from Ehrlich ascites cells. Antisense riboprobes extending from the 5' ends of the pDR100 and pDR101 cDNAs to the SmaI restriction site within the mLIF open reading frame were produced by <u>in vitro</u> transcription of these clones. The protection patterns obtained with these riboprobes were compared with the known pattern generated by the 5' riboprobe derived from pDR2.

The results and interpretation of this experiment are shown in Figure 6. Each riboprobe protected a fragment of 345 nucleotides, the distance between the 5' end of mLIF exon 2 and the SmaI restriction site. This band (c) therefore corresponds to transcript M for probes pDR2 and

pDR100, and transcript D for probe pDR101. Probes pDR100 and PDR101 both protected fragments larger than those protected by probe pDR2 (345 and 369 nucleotides). This establishes that the sequences upstream of the ATG in these RACE clones are present in the cellular D and M transcripts respectively, and confirms the identification of clone pDR100 as a cDNA encoding the diffusible form of DIA/LIF. The presence of two protected bands larger than 369 nucleotides suggests that the promoters directing expression of the two DIA/LIF transcripts each contain two RNA initiation sites. The deduced positions of these sites are indicated in Figure 5.

REFERENCES

Bradley, A. et al, (1984) Nature 309:225-256.

Edwards, D. et al, (1985) Mol. Cell. Biol. 5:3280-3288.

Edwards, D. et al, (1987) EMBO J. 6:1899-1904.

Evans, M.J., and Kaufman, M. (1983) Cancer Surveys 2:185-206.

Frohman, et al, (1988) Proc.Natl.AcadSci.USA 85 (23). 1988. 8998-9002.

Gearing, D.F., et al EMBO (1987) J. 6:3995-4002.

Gearing, D.F., et al (1988) Nuc. Acids. Res. 16:9857.

Gospodarawicz, et al J. (1984) Cell Biol. 99:947-961.

Gough, N.M., et al (1988) Proc. Natl. Acad. Sci. U.S.A. 85:2623-7

Krieg, P., and Melton, D.A. (1987) Methods in Enzymology 155:397-415.

Maniatis, T., et al (1982) Molecular Cloning: a laboratory Manual. Cold Spring Harbour press.

Martin, G. (1981) Proc. Natn. Acad. Sci. U.S.A. 78:7634-7638.

Martin G.R., and Evans, M (1975) Proc. Natn. Acad. Sci. U.S.A. 72:1441-1445.

Moreau, J.F., et al (1988) Nature 326:201-204.

Rheinwald, J., and Green, H., (1975) Cell 6:331-344.

Smith, A.G., et al (1988) Nature 336:688-690.

Smith, A.G., and Hooper, M.L., (1987) Dev. Biol. 121:1-9.

CLAIMS

1. A substantially pure Leukaemia Inhibitory Factor (LIF) which is capable of associating with a mammalian extracellular matrix.

2. A LIF according to claim 1 wherein the N-terminal sequence comprises:

NH_2-U-arg-X-arg-

where U is the residue of a peptide of 20 or fewer amino acids, or a covalent bond, and X is any amino acid.

3. A LIF according to claim 2 wherein U represents the residue of methioneine and X is cys or gly.

4. A LIF according to either of claims 2 or 3 wherein the N terminal sequence comprises:

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U and X are as defined in claim 2, and substantially identical variants thereof.

5. A recombinant DNA encoding the LIF of any one of claims 1 to 4.

6. A recombinant replicable expression vector comprising the DNA of claim 5.

7. A host cell transformed or transfected with the vector of claim 6.

8. A polyclonal or monoclonal antibody directed against the N-terminal of a LIF according to any one of claims 1 to 4, or produced by a host cell according to claim 7.

9. A pharmaceutical composition comprising a LIF according to any one of claims 1 to 4 or any antibody according to claim 8 in association with a pharacologically acceptable carrier or diluent.

10. A recombinant protein or peptide which comprises, at its N-terminus, the sequence:

NH_2-U-arg-X-arg-

where U is the residue of a peptide of 20 or fewer amino acids, or a covalent bond, and X is any amino acid.

11. A recombinant protein or peptide according to either of claims 9 or 10 wherein the N terminal sequence comprises:

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U and X are as defined in claim 8, and substantially identical variants thereof.

12. A recombinant protein or peptide according to claim 10 or 11 which is a hormone, growth factor or cytokine.

13. A recombinant DNA encoding a protein or peptide according to any one of claims 9 to 12.

14. A recombinant replicable expression vector comprising the DNA of claim 13.

15. A host cell transformed or transfected with the vector of claim 14.

16. A polyclonal or monoclonal antibody directed against the N-terminal of a protein or peptide according to any one of claims 9 to 14, or produced by a host cell according to claim 15.

17. A peptide of the formula:

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-COOH

where U is the residue of a peptide of 20 or fewer amino acids or a covalent bond and X is any amino acid, and substantially identical variants thereof.

18. A method of causing a protein or peptide to become attached to the ECM of a mammalian cell comprising contacting the ECM with a protein or peptide having an N-terminus of the sequence:

NH_2-U-arg-X-arg-

where U is the residue of a peptide of 20 or fewer amino acids or a covalent bond and X is any amino acid.

19. A method according to claim 18 wherein the N-terminal of the protein or peptide comprises:

NH_2-U-arg-X-arg-ile-val-pro-leu-leu-

where U and X are as defined in claim 18, and substantially identical variants thereof.

20. The use of LIF according to any one of claims 1 to 4 or an antibody according to claim 8 in method of treatment of the human or animal body by therapy or diagnosis.

21. A method of therapy of conditions associated with abnormal and/or aberrant expression of LIF in a mammal, including man, which will benefit from an alteration in naturally occuring levels of LIF which comprises administering to a recipient in need of such therapy an effective amount of a LIF according to claim 1 or an antibody according to claim 8.

Fig.1.

EVIDENCE FOR A MATRIX-ASSOCIATED FORM OF LIF/DIA.

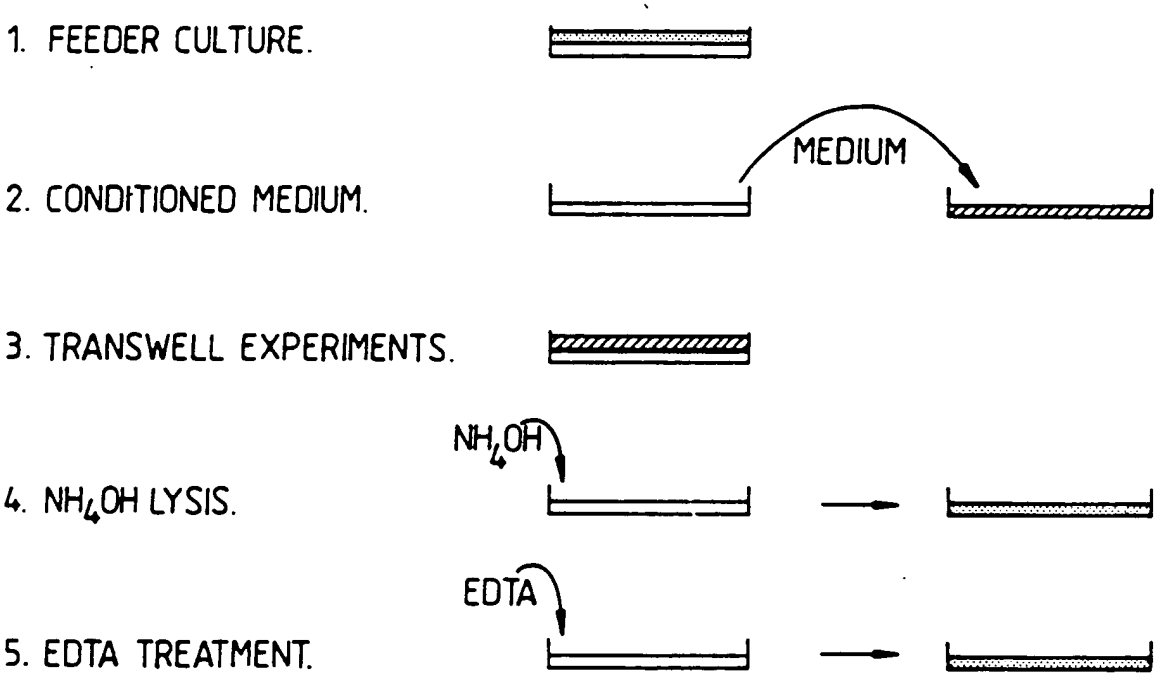


Fig. 2.

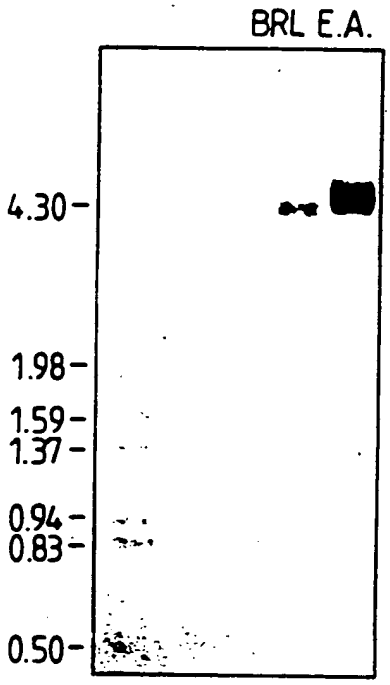


Fig. 3A.

Fig. 3B.

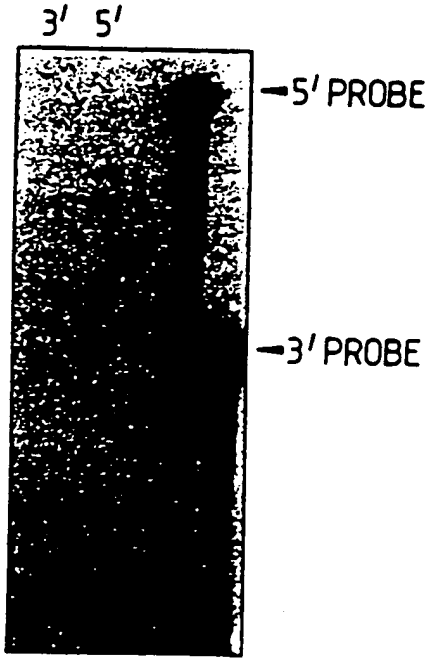


Fig.4A.

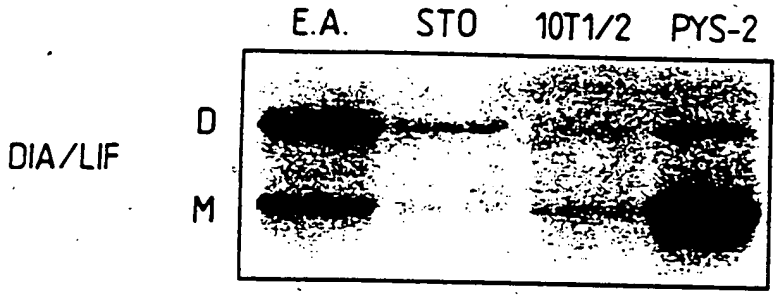


Fig.4B.

DIA/LIF ACTIVITY	E.A.	STO	10T1/2	PYS-2
MEDIUM	++	+/-	-	n.a.
MATRIX	n.a.	+	+	+

Fig.5.

ALTERNATIVE FORMS OF DIA/LIF

CTAGTCCCTGGAAAGCTGTGATTGGCGCGAGATGAGATGCAGGATT MATRIX

GGAGTCCAGCCCATAATGAAGGTCTTGGCCGCAGGGATT DIFFUSIBLE

M R C R I V P L L L L V L H MATRIX

M K V L A A G I V P L L L L V L H DIFFUSIBLE

Fig:6A.

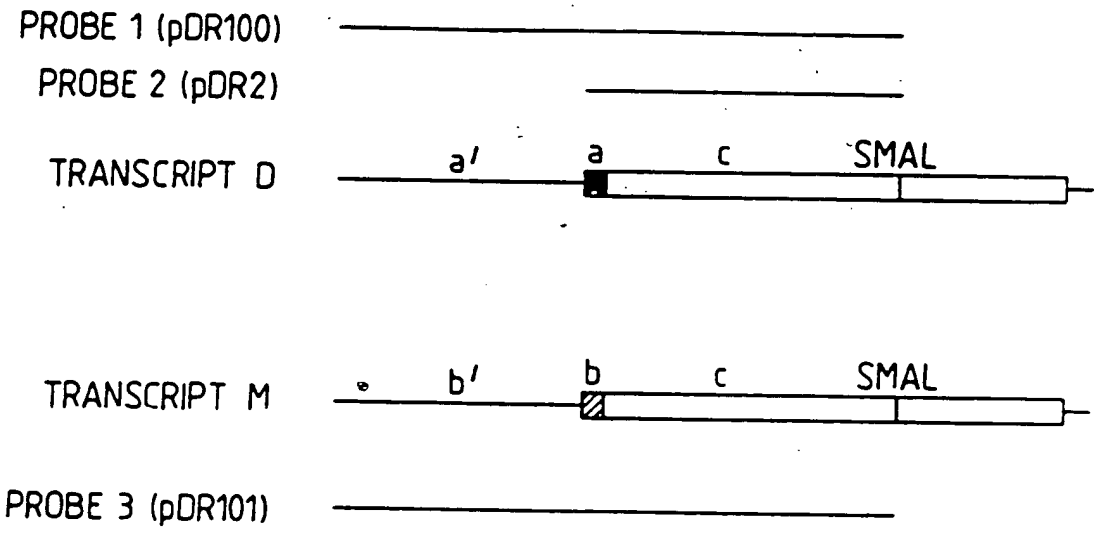
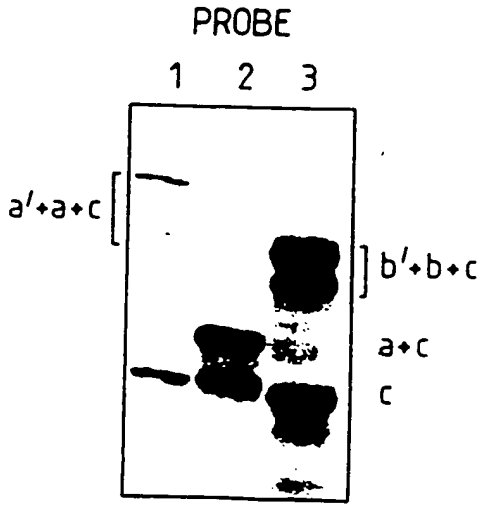

PROBE 1 (pDR100)
PROBE 2 (pDR2)
TRANSCRIPT D
a'
a
c
SMAL
TRANSCRIPT M
b'
b
c
SMAL
PROBE 3 (pDR101)

Fig.6B.


PROBE
1 2 3
a'+a+c
b'+b+c
a+c
c

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 91/00334

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 07 K 13/00, 7/06, C 12 N 15/62, C 12 P 21/08, C 12 N 15/24, A 61 K 37/02, C 07 K 3/08, C 12 N 1/21, A 61 K 39/395

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N, C 07 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X	Cell, volume 62, 21 September 1990, Cell Press, (Cambridge, NA, GB), P.D. Rathjen et al.: "Differentiation inhibiting activity is produced in matrix-associated and diffusible forms that are generated by alternate promoter usage", pages 1105-1114 see the whole document	1-7,10-15, 17-19
A	The EMBO Journal, volume 6, no. 13, 20 December 1987, (Heidelberg, DE), D.P. Gearing et al. "Molecular cloning and expression of cDNA encoding a murine myeloid leukaemia inhibitory factor (LIF)", pages 3995-4002 see the whole document cited in the applcation	1-9

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
27th May 1991	08.08.91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	Mme. M. van der Drift

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	Proceedings of The national Academy of Sciences of the USA, volume 85, no. 8, April 1988, Cell Biology, (Washington, DC, US), N.M. Gough et al.: "Molecular cloning and expression of the human homologue of the murine gene encoding myeloid leukemia-inhibitory factor", pages 2623-2627 see the whole document cited in the application	1-9
X	Proceedings of The National Academy of Sciences, volume 84, January 1987, (Washington, DC, US), D.R. Joseph et al.: "Rat androgen-binding protein: Evidence for identical subunits and amino acid sequence homology with human sex hormone-binding globulin", pages 339-343 see figure 4	10
X	The Journal of Biological Chemistry, volume 263, no. 14, 15 May 1988, The American Society for Biochemistry and Molecular Biology, Inc., (US), T. Pikkarainen et al.: "Human laminin B2 chain", pages 6751-6758 see figure 3	10

FURTHER INFORMATI N CONTINUED FROM THE SECOND SHEET

V. ☒ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☒ Claim numbers 20, 21 because they relate to subject matter not required to be searched by this Authority, namely:

Pls. see Rule 39.1(iv) - PCT:
Methods for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods.

2. ☐ Claim numbers _____, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers _____, because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

The International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.